Publ. Mat. **64** (2020), 255–281
DOI: 10.5565/PUBLMAT6412011

ON A BINARY SYSTEM OF PRENDIVILLE: THE CUBIC CASE

Shaoming Guo

Abstract: We prove sharp decoupling inequalities for a class of two dimensional non-degenerate surfaces in $\mathbb{R}^5$, introduced by Prendiville [**13**]. As a consequence, we obtain sharp bounds on the number of integer solutions of the Diophantine systems associated with these surfaces.

2010 Mathematics Subject Classification: Primary: 11L07; Secondary: 42A45.

Key words: decoupling, transversality, Brascamp–Lieb inequalities, Diophantine systems.

1. Introduction

Let $\Phi(t,s)$ be a homogeneous polynomial of degree three. Consider the two dimensional surface

$$\mathcal{S} = \{(t, s, \Phi_t(t,s), \Phi_s(t,s), \Phi(t,s)) : (t,s) \in [0,1]^2\}. \tag{1.1}$$

We say that Φ is non-degenerate if it can not be written as $(\mu t + \nu s)^3$ for any $\mu, \nu \in \mathbb{R}$. This is the same as saying that, if we write $\Phi(t,s) = at^3 + bt^2 s + cts^2 + ds^3$, then the matrix

$$\begin{pmatrix} 3a & 2b & c \\ b & 2c & 3d \end{pmatrix} \tag{1.2}$$

has rank two. This will be our assumption throughout the present paper.

Consider the following system of Diophantine equations

$$\begin{cases} x_1 + x_2 + \cdots + x_r = x_{r+1} + x_{r+2} + \cdots + x_{2r}, \\ y_1 + y_2 + \cdots + y_r = y_{r+1} + y_{r+2} + \cdots + y_{2r}, \\ \Phi_t(x_1,y_1) + \cdots + \Phi_t(x_r,y_r) = \Phi_t(x_{r+1},y_{r+1}) + \cdots + \Phi_t(x_{2r},y_{2r}), \\ \Phi_s(x_1,y_1) + \cdots + \Phi_s(x_r,y_r) = \Phi_s(x_{r+1},y_{r+1}) + \cdots + \Phi_s(x_{2r},y_{2r}), \\ \Phi(x_1,y_1) + \cdots + \Phi(x_r,y_r) = \Phi(x_{r+1},y_{r+1}) + \cdots + \Phi(x_{2r},y_{2r}). \end{cases} \tag{1.3}$$

Here r is a positive integer and $x_i, y_i \in \mathbb{N}$ for each $1 \le i \le 2r$. For a large integer N, we let $J_r(N)$ denote the number of integer solutions $(x_1, \ldots, x_{2r}, y_1, \ldots, y_{2r})$ of the system (1.3) with $0 \le x_i, y_i \le N$ for each $1 \le i \le 2r$. We prove

Theorem 1.1. *For each $r \ge 1$ and each $\epsilon > 0$, we have*

$$J_r(N) \lesssim_{r,\epsilon} N^{2r+\epsilon} + N^{4r-9+\epsilon}. \tag{1.4}$$

Here the implicit constant depends only on r and ϵ. Moreover, up to the arbitrarily small factor ϵ, the exponents of N are sharp.

The lower bounds have been calculated by Parsell, Prendiville, and Wooley [**12**]. Our focus is to obtain the upper bounds (1.4). This will be done via proving a sharp decoupling inequality.

For a measurable set $R \subset [0,1]^2$ and a measurable function $g\colon R \to \mathbb{C}$, define the extension operator associated with $\mathcal{S}$ by

$$E_R g(x) = \int_R g(t,s) e^{itx_1 + isx_2 + i\Phi_t(t,s)x_3 + i\Phi_s(t,s)x_4 + i\Phi(t,s)x_5} \, dt\, ds. \tag{1.5}$$

Here $x = (x_1, \ldots, x_5)$. For a ball $B = B(c,R) \subset \mathbb{R}^5$ with center c and radius R, we use the weight

$$w_B(x) = \left(1 + \frac{\|x - c\|}{R}\right)^{-C}, \tag{1.6}$$

where C is a large enough constant whose value will not be specified. For each $2 \le q \le p$ and $0 < \delta < 1$, let $B_{p,q}(\delta)$ be the smallest constant such that

$$\|E_{[0,1]^2} g\|_{L^p(w_B)} \le B_{p,q}(\delta) \Bigg(\sum_{\substack{\Delta\colon \text{square in } [0,1]^2 \\ l(\Delta) = \delta}} \|E_\Delta g\|^q_{L^p(w_B)} \Bigg)^{1/q} \tag{1.7}$$

holds for each ball $B \subset \mathbb{R}^5$ of radius δ^{-3}. Inequalities of this type are referred to as $l^q L^p$ decouplings.

Theorem 1.2. *We have*

$$B_{9,9}(\delta) \lesssim \left(\frac{1}{\delta}\right)^{2(\frac{1}{2} - \frac{1}{9}) + \epsilon}, \tag{1.8}$$

for each $\epsilon > 0$ and $0 < \delta \le 1$.

Via a standard reduction (see for instance [**5**, Section 2]), Theorem 1.1 follows from Theorem 1.2. In this reduction, one takes the function g in (1.7) to be a linear combination of Dirac functions, hence $E_{[0,1]^2} g$ becomes an exponential sum. From a good exponential sum estimate, one applies periodicity to conclude a good upper bound on the number of integer solutions of the corresponding Diophantine system, which is (1.3) in our case.

The system (1.3) is the cubic case of a system considered by Prendiville [**13**]. The way that the surface (1.1) and the Diophantine system (1.3) are formulated is slightly different from those in Prendiville [**13**]. There, the surface (1.1) is replaced by

$$\mathcal{S}' = \{(\Phi_{tt}(t,s), \Phi_{ts}(t,s), \Phi_{ss}(t,s), \Phi_t(t,s), \Phi_s(t,s), \Phi(t,s)) : (t,s) \in [0,1]^2\}. \tag{1.9}$$

That is, the surface $\mathcal{S}'$ is obtained by taking successive partial derivatives of the seed polynomial Φ. However, under the non-degeneracy condition that the matrix (1.2) has rank two, we observe that the vector space $[\Phi_{tt}, \Phi_{ts}, \Phi_{ss}]$ is always the same as $[t, s]$. Hence the system of Diophantine equations associated with the surface $\mathcal{S}'$ is always equivalent to that associated with the surface $\mathcal{S}$, in the sense that they admit the same number of integer solutions.

To obtain a system analogous to (1.3) of higher degrees, one takes a seed polynomial $\Phi(t,s)$ of degree $k \geq 3$, extracts all the partial derivatives

$$\frac{\partial^{i_1+i_2}\Phi(t,s)}{\partial t^{i_1}\partial s^{i_2}} \quad (i_1 \geq 0,\, i_2 \geq 0), \tag{1.10}$$

and forms a Diophantine system by using all these partial derivatives. If we take $\Phi(t,s)$ to be the monomial $t^{k_1}s^{k_2}$ with $k_1 \geq k_2 \geq 1$, then we recover the so-called simple binary systems

$$x_1^{i_1}y_1^{i_2} + \cdots + x_r^{i_1}y_r^{i_2} = x_{r+1}^{i_1}y_{r+1}^{i_2} + \cdots + x_{2r}^{i_1}y_{2r}^{i_2}, \quad \text{with } i_1 \leq k_1,\, i_2 \leq k_2, \text{ and } (i_1,i_2) \neq (0,0), \tag{1.11}$$

which appeared in recent work in quantitative arithmetic geometry ([**14**, Section 4.15] and [**15**]). Notice that if we take Φ to be a polynomial of degree k that depends only on one variable, then we recover the Vinogradov system

$$x_1^i + \cdots + x_r^i = x_{r+1}^i + \cdots + x_{2r}^i, \text{ with } 1 \leq i \leq k. \tag{1.12}$$

All the systems mentioned above fall into the framework of translation-dilation invariant systems, which are intensively studied in [**12**]. In our setting, this is reflected in the validity of the parabolic rescaling lemma (Lemma 3.1).

Parsell, Prendiville, and Wooley ([**12**]) proved (1.4) for $r \geq 21$, using the method of efficient congruencing. In the current paper we prove it for all $r \geq 1$, using the decoupling theory developed in [**3**] and [**7**].

When intending to generalise our proof to the above binary systems ((1.10) or (1.11)) of degrees higher than three, one encounters enormous difficulties. In comparison, the efficient congruencing method still provides bounds that are almost optimal. We refer to [**12**] for the precise statement of the corresponding results.

Let us mention a further application of the result in Theorem 1.2. This application has been worked out carefully in [**13**], [**12**], and [**11**], hence we mention it briefly. Let Φ be as above, a homogeneous polynomial of degree three that is non-degenerate. Take $r \in N$. Let $c_1, c_2, \dots, c_r$ with $c_1 + c_2 + \cdots + c_r = 0$ be a "non-singular" choice of coefficients for Φ (see [**13**, Definition 1.1]). Consider the equation

$$c_1\Phi(x_1, y_1) + c_2\Phi(x_2, y_2) + \cdots + c_r\Phi(x_r, y_r) = 0. \tag{1.13}$$

The solution $\{(x_1, y_1), \dots, (x_r, y_r)\}$ to the above equation is called diagonal if they all lie on a line in the plane. Take a large number $N \in \mathbb{N}$. Let $A \subset [0, N]^2$ be a set which contains only diagonal solutions to the equation (1.13). Then a result in [**13**] (further improved in [**12**]) states that

$$|A| \ll N^2(\log\log N)^{-1/(s-1)}, \tag{1.14}$$

for s bigger than a certain threshold. The validity of the estimate (1.8) will further lower down this threshold. We refer the interested reader to [**13**] and [**11**] for the details.

In the end, we mention some novelties of our proof and explain briefly the potential difficulties that appear when trying to adapt our argument to binary systems of higher degrees.

In decoupling theory, various Brascamp–Lieb inequalities (see (2.2)) play fundamental roles. In order to apply these inequalities, one needs to check a transversality condition (see (2.3)). When the dimensions and co-dimensions of the surfaces under consideration get higher and higher, checking these transversality conditions will become more and more difficult. In the current paper we are dealing with a two dimensional surface in $\mathbb{R}^5$. To check (2.3), we further develop the idea introduced in [**6**], where a specific two dimensional surface in $\mathbb{R}^9$ is considered. As currently we are dealing with a class of surfaces, certain algebraic structures need to be explored. For instance, see Subsection 2.2, and in particular Lemma 2.3.

Notation. Throughout the paper we will write $A \lesssim_v B$ to denote the fact that $A \le CB$ for a certain implicit constant C that depends on the parameter v. Typically, this parameter is either ϵ or K. The implicit constant will never depend on the scale δ, on the balls we integrate over, or on the function g. It will however most of the times depend on the Lebesgue index p.

We will denote by B_R an arbitrary ball of radius R. We use the following notation for averaged integrals

$$\|F\|_{L^p_{\#}(w_B)} = \left(\frac{1}{|B|}\int |F|^p w_B\right)^{1/p}.$$

$|A|$ will refer to either the cardinality of A if A is finite, or to its Lebesgue measure if A has positive measure.

Acknowledgements. The author thanks Ciprian Demeter for reading this paper and giving several very useful suggestions. The author also thanks Sean Prendiville for discussions on the applications of our main result. Last but not least, the author also thanks the referee for many helpful suggestions that significantly improved the exposition of the paper.

2. Brascamp–Lieb inequalities and ball-inflation lemmas

Let m be a positive integer. For $1 \le j \le m$, let V_j be a d-dimensional linear subspace of $\mathbb{R}^n$. Let also $\pi_j \colon \mathbb{R}^n \to V_j$ denote the orthogonal projection onto V_j. Define

$$\Lambda(f_1, f_2, \dots, f_m) = \int_{\mathbb{R}^n} \prod_{j=1}^{m} f_j(\pi_j(x))\, dx, \tag{2.1}$$

for $f_j \colon V_j \to \mathbb{C}$. We recall the following theorem due to Bennett, Carbery, Christ, and Tao [**2**].

Theorem 2.1 ([**2**]). *Given $p \ge 1$, the estimate*

$$|\Lambda(f_1, f_2, \dots, f_m)| \lesssim \prod_{j=1}^{m} \|f_j\|_p \tag{2.2}$$

holds if and only if $np = dm$ and the following Brascamp–Lieb transversality condition is satisfied

$$\dim(V) \le \frac{1}{p}\sum_{j=1}^{m} \dim(\pi_j(V)), \quad \textit{for each linear subspace } V \subset \mathbb{R}^n. \tag{2.3}$$

An equivalent formulation of the estimate (2.2) is

$$\left\| \left(\prod_{j=1}^{m} g_j \circ \pi_j\right)^{1/m} \right\|_q \lesssim \left(\prod_{j=1}^{m} \|g_j\|_2\right)^{1/m}, \tag{2.4}$$

with $q = \frac{2n}{d}$. The restriction that $p \geq 1$ becomes $dm \geq n$. In our proof, m will always be a large constant, hence this condition is always satisfied. The transversality condition (2.3) becomes

$$\dim(V) \leq \frac{n}{dm} \sum_{j=1}^{m} \dim(\pi_j(V)), \text{ for each subspace } V \subset \mathbb{R}^n. \tag{2.5}$$

Let us be more precise about the parameters in (2.5). We will take $n = 5$ as our surface $\mathcal{S}$ lives in $\mathbb{R}^5$. The degree m of multi-linearity will be chosen to be a large number. Our proof will make use of two different values of the parameter d: First of all, we will use $d = 2$, which corresponds to the fact that the surface $\mathcal{S}$ is two-dimensional; secondly, we also need to use $d = 4$, as at certain stage of the proof we will view $\mathcal{S}$ as a four-dimensional surface in $\mathbb{R}^5$. For instance, see Lemma 2.5 in Subsection 2.3.

Recall that the surface we are looking at is $(t,s,\Phi_t(t,s),\Phi_s(t,s),\Phi(t,s))$. Its tangent space is spanned by

$$\overrightarrow{n_1} = (1, 0, \Phi_{tt}, \Phi_{st}, \Phi_t) \text{ and } \overrightarrow{n_2} = (0, 1, \Phi_{ts}, \Phi_{ss}, \Phi_s). \tag{2.6}$$

Moreover, we denote

$$\overrightarrow{n_3} = (0, 0, 1, 0, t) \text{ and } \overrightarrow{n_4} = (0, 0, 0, 1, s). \tag{2.7}$$

We will see from the following Lemma 2.3 that these two vectors span the "second order tangent space". At a point $\xi \in [0, 1]^2$, let $V_\xi^{(1)}$ be the linear space spanned by $\overrightarrow{n_1}(\xi)$ and $\overrightarrow{n_2}(\xi)$ given in (2.6). Let $V_\xi^{(2)}$ be the linear space spanned by $\overrightarrow{n_1}$, $\overrightarrow{n_2}$, $\overrightarrow{n_3}$, $\overrightarrow{n_4}$ at the point ξ.

Let $K \in \mathbb{N}$ be a large number. It will be sent to infinity at the end of our proof. A K-square is defined to be a closed square of length $1/K$ inside $[0, 1]^2$. The collection of all dyadic K-squares will be denoted by Col_K.

Proposition 2.2. *Take $\Lambda \in \mathbb{N}$. Denote $m = \Lambda K$. Let $R_1, R_2, \ldots, R_m$ be distinct K-squares from Col_K. For each $1 \leq i \leq m$, choose one point $\xi_i \in R_i$. If we choose Λ sufficiently large, independently of any parameter, then the transversality condition* (2.5) *with* $(d, n) = (2, 5)$ *(respectively* $(4, 5)$*) is satisfied for the collection of spaces $\{V_{\xi_j}^{(1)}\}_{j=1}^m$ (respectively $\{V_{\xi_j}^{(2)}\}_{j=1}^m$).*

We will prove Proposition 2.2 in the following two subsections. How to check the Brascamp–Lieb transversality condition (2.5) seems to be a big obstacle in obtaining new decoupling inequalities associated with surfaces of high co-dimensions. For instance, see [**6**], where a particular two dimensional surface in $\mathbb{R}^9$ is considered. The forthcoming argument that corresponds to the case $d = 4$ in Subsection 2.2 further develops the idea introduced in [**6**]. From our argument, in particular Lemma 2.3, it will become clear that more algebraic structures need to be understood in order to push our current results to homogeneous polynomials of degrees higher than three.

2.1. Proof of Proposition 2.2: The case $d = 2$. In this subsection we prove the first part of Proposition 2.2. Let $\pi_\xi^{(1)}(V)$ denote the projection of the space V on $V_\xi^{(1)}$. We will show that

$$\dim(V) \leq \frac{5}{2}\dim(\pi_\xi^{(1)}(V)) \text{ almost surely in } \xi. \tag{2.8}$$

Assume that $\dim(V) < 5$, as otherwise the statement is trivial. Assume that the linear space V is spanned by $\overrightarrow{v_1}, \ldots, \overrightarrow{v}_{\dim(V)}$. By the rank nullity theorem, the dimension of $\pi_\xi^{(1)}(V)$ is equal to the rank of the matrix

$$\left[\langle \overrightarrow{n_i}, \overrightarrow{v_j} \rangle\right]_{\substack{1\leq i\leq 2;\\ 1\leq j\leq \dim(V)}}. \tag{2.9}$$

Therefore, proving (2.8) is equivalent to proving that the matrix (2.9) has a minor of order

$$\left[\frac{2\cdot\dim(V)}{5}\right] + 1, \tag{2.10}$$

whose determinant does not vanish constantly. Here, for $a > 0$, the symbol $[a]$ refers to the biggest integer that is smaller than or equal to a.

Let us assume (2.8) for a moment and see how it implies the transversality condition (2.5). First of all, if we define an exceptional set

$$E_V := \left\{\xi \in [0,1]^2 : \dim(V) > \frac{5}{2}\dim(\pi_\xi^{(1)}(V))\right\}, \tag{2.11}$$

then (2.8) implies that E_V lies inside the zero set of a polynomial of degree less than 10. This is because every entry in the matrix (2.9) is a polynomial of order at most two. Hence, the determinant of a minor of (2.9) of an order given by (2.10) is a polynomial of order at most 4, which is smaller than 10.

However, Wongkew's lemma ([**16**]) says that the $\frac{10}{K}$-neighbourhood of the zero set of such a polynomial will intersect at most CK squares in Col_K for some large constant C. In general, this constant C in Wongkew's lemma depends on the dimension and the degree of a polynomial. As we are working on polynomials of two variables and degrees bounded by 10, the constant C here can be chosen as a universal constant. The desired transversality condition (2.5) follows immediately if we choose $\Lambda = 100C$.

Case $\dim(V) = 1$ *or* 2. The desired estimate (2.8) is reduced to

$$\dim(\pi_\xi^{(1)}(V)) = 1 \text{ almost surely.} \tag{2.12}$$

Suppose $V = \text{span}\{\overrightarrow{u}\}$ with $\overrightarrow{u} = (u_1, u_2, u_3, u_4, u_5)$. Then (2.12) is equivalent to

$$(\overrightarrow{u} \cdot \overrightarrow{n_1}, \overrightarrow{u} \cdot \overrightarrow{n_2}) \not\equiv (0, 0). \tag{2.13}$$

We argue by contradiction. Suppose $(\overrightarrow{u} \cdot \overrightarrow{n_1}, \overrightarrow{u} \cdot \overrightarrow{n_2}) = (0, 0)$ for every $\xi \in [0, 1]^2$. By checking the constant terms in the polynomials $\overrightarrow{u} \cdot \overrightarrow{n_1}$ and $\overrightarrow{u} \cdot \overrightarrow{n_2}$, we obtain $u_1 = u_2 = 0$. By checking the highest order terms, we obtain $u_5 = 0$. These two facts further imply that the cross product

$$(\Phi_{tt}, \Phi_{st}) \times (\Phi_{ts}, \Phi_{ss}) \tag{2.14}$$

is constantly zero. However, by a direct calculation, this contradicts the assumption that the polynomial Φ is non-degenerate.

Case $\dim(V) = 3$ *or* 4. We need to show that $\dim(\pi_\xi^{(1)}(V)) \geq 2$ almost surely. This is done via a direct calculation. Clearly the case $\dim(V) = 3$ is more difficult. Suppose $V = \text{span}\{\overrightarrow{u}, \overrightarrow{v}, \overrightarrow{w}\}$. Then the dimension of $\pi_\xi^{(1)}(V)$ is equal to the rank of the matrix

$$\begin{pmatrix} \overrightarrow{u} \cdot \overrightarrow{n_1} & \overrightarrow{v} \cdot \overrightarrow{n_1} & \overrightarrow{w} \cdot \overrightarrow{n_1} \\ \overrightarrow{u} \cdot \overrightarrow{n_2} & \overrightarrow{v} \cdot \overrightarrow{n_2} & \overrightarrow{w} \cdot \overrightarrow{n_2} \end{pmatrix}. \tag{2.15}$$

We argue by contradiction and suppose that the determinants of all the two by two minors vanish constantly. We look at the two by two minor formed by the first two columns. The determinant of the matrix

$$\begin{pmatrix} u_1 + u_3\Phi_{tt} + u_4\Phi_{st} + u_5\Phi_t & v_1 + v_3\Phi_{tt} + v_4\Phi_{st} + v_5\Phi_t \\ u_2 + u_3\Phi_{ts} + u_4\Phi_{ss} + u_5\Phi_s & v_2 + v_3\Phi_{ts} + v_4\Phi_{ss} + v_5\Phi_s \end{pmatrix} \tag{2.16}$$

vanishes constantly. Denote

$$d_{i,j} := \det \begin{pmatrix} u_i & u_j \\ v_i & v_j \end{pmatrix}. \tag{2.17}$$

We first look at the third order term, that is,

$$\begin{aligned} & d_{5,4}\Phi_t\Phi_{ss} + d_{5,3}\Phi_t\Phi_{ts} + d_{3,5}\Phi_{tt}\Phi_s + d_{4,5}\Phi_{st}\Phi_s \\ & \qquad = \left(d_{3,5}\frac{\partial}{\partial_t}\left(\frac{\Phi_t}{\Phi_s}\right) + d_{4,5}\frac{\partial}{\partial_s}\left(\frac{\Phi_t}{\Phi_s}\right)\right)\Phi_s^2 \equiv 0. \end{aligned} \tag{2.18}$$

This further implies $d_{3,5} = d_{4,5} = 0$. Moreover, we know $d_{1,2} = 0$ by checking the constant term of the determinant of the matrix (2.16). This further implies that

$$(u_5, v_5, w_5) = (0, 0, 0), \tag{2.19}$$

as otherwise we would derive a contradiction that $(\overrightarrow{u}, \overrightarrow{v}, \overrightarrow{w})$, when viewed as a matrix of order 3×5, has rank two or smaller.

Substitute the identity (2.19) into (2.16) and look at the second order term of the determinant of (2.16). We obtain

$$d_{3,4}\Phi_{tt}\Phi_{ss} - d_{3,4}\Phi_{st}\Phi_{st} \equiv 0. \tag{2.20}$$

By the non-degeneracy assumption on Φ, we obtain that $d_{3,4} = 0$. This, together with (2.19) and $d_{1,2} = 0$, implies that the 3×5 matrix (u, v, w) has rank two or smaller, which is a contradiction.

2.2. Proof of Proposition 2.2: The case $d = 4$. We let $\pi_\xi^{(2)}(V)$ denote the projection of the space V on $V_\xi^{(2)}$. We need to show that

$$\dim(V) \leq \frac{5}{4}\dim(\pi_\xi^{(2)}(V)) \text{ almost surely.} \tag{2.21}$$

This amounts to calculating the dimension of

$$\{(\overrightarrow{u} \cdot \overrightarrow{n_1}, \overrightarrow{u} \cdot \overrightarrow{n_2}, \overrightarrow{u} \cdot \overrightarrow{n_3}, \overrightarrow{u} \cdot \overrightarrow{n_4}) : \overrightarrow{u} \in V\}. \tag{2.22}$$

Following [**6**], we define linear spaces

$$S_1 = [t, s],\ S_2 = [\Phi_t(t, s), \Phi_s(t, s)],\ \text{and } S_3 = [\Phi(t, s)]. \tag{2.23}$$

We need the following version of Taylor's formula.

Lemma 2.3. *If $f \in S_3$, then*

$$\Delta f(t, s) \approx f_t(t, s)\Delta t + f_s(t, s)\Delta s + t \cdot f_t(\Delta t, \Delta s) + s \cdot f_s(\Delta t, \Delta s). \tag{2.24}$$

Here $\Delta f(t, s) = f(t + \Delta t, s + \Delta s) - f(t, s)$. The error produced by the approximate identity is a third order homogeneous polynomial in Δt and Δs.

Proof: By linearity, it suffices to consider $f(t, s) = \Phi(t, s)$. We calculate $\Phi(t + \Delta t, s + \Delta s) - \Phi(t, s)$ and view it as a homogeneous polynomial of four variables t, s, Δt, and Δs. First, we collect the linear terms with respect to Δt and Δs. By the first order Taylor expansion, they are given by $\Phi_t(t, s)\Delta t + \Phi_s(t, s)\Delta s$, which are the former two terms on the right hand side of (2.24). Next, we collect the quadratic terms with respect to Δt and Δs. These terms must be linear in the variables t and s. We apply the first order Taylor expansion again, with the roles of (t,s) and $(\Delta t,\Delta s)$ exchanged, and obtain $t\Phi_t(\Delta t,\Delta s)+s\Phi_s(\Delta t,\Delta s)$, which gives the latter two terms in (2.24). $\square$

This lemma can be written in the following equivalent way.

$$f(t,s)-f(t_0,s_0)\approx f_t(t_0,s_0)(t-t_0)+f_s(t_0,s_0)(s-s_0) + t_0\cdot f_t(t-t_0,s-s_0)+s_0\cdot f_s(t-t_0,s-s_0). \tag{2.25}$$

According to this formula, let us consider

$$f(t,s)=u_1t+u_2s+u_3\Phi_t(t,s)+u_4\Phi_s(t,s)+u_5\Phi(t,s). \tag{2.26}$$

At each point $\xi=(t_0,s_0)$, denote $\Delta t=t-t_0$ and $\Delta s=s-s_0$. We define

$$(P_\xi f)(t,s)=f(\xi)+f_t(\xi)\cdot\Delta t+f_s(\xi)\cdot\Delta s + (u_3+u_5t_0)\Phi_t(\Delta t,\Delta s)+(u_4+u_5s_0)\Phi_s(\Delta t,\Delta s). \tag{2.27}$$

Here we observe that

$$P_\xi f=f \text{ for } f\in S_1\oplus S_2. \tag{2.28}$$

We further define the canonical projection $\pi_{S_1\oplus S_2}$ onto the space $S_1\oplus S_2$. Hence

$$\begin{aligned}(\pi_{S_1\oplus S_2}P_\xi f)(t,s)=&(f_t(\xi)+(u_3+u_5t_0)\Phi_{tt}(-t_0,-s_0)\\ &\quad+(u_4+u_5s_0)\Phi_{st}(-t_0,-s_0))t\\ &+(f_s(\xi)+(u_3+u_5t_0)\Phi_{ts}(-t_0,-s_0)\\ &\quad+(u_4+u_5s_0)\Phi_{ss}(-t_0,-s_0))s\\ &+(u_3+u_5t_0)\Phi_t(t,s)+(u_4+u_5s_0)\Phi_s(t,s)\\ =&(u_1+u_5\Phi_t(\xi)-u_5t_0\Phi_{tt}(\xi)-u_5s_0\Phi_{st}(\xi))t\\ &\quad+(u_2+u_5\Phi_s(\xi)-u_5t_0\Phi_{ts}(\xi)-u_5s_0\Phi_{ss}(\xi))s\\ &+(u_3+u_5t_0)\Phi_t(t,s)+(u_4+u_5s_0)\Phi_s(t,s).\end{aligned} \tag{2.29}$$

We can write

$$\begin{aligned}\pi_{S_1\oplus S_2}P_\xi f=(&u_1+u_5\Phi_t(\xi)-u_5t_0\Phi_{tt}(\xi)-u_5s_0\Phi_{st}(\xi),\\ &u_2+u_5\Phi_s(\xi)-u_5t_0\Phi_{ts}(\xi)-u_5s_0\Phi_{ss}(\xi),\\ &u_3+u_5t_0,u_4+u_5s_0).\end{aligned} \tag{2.30}$$

Recall the choice of the function f in (2.26). Let us compare the vector in (2.30) with the vector in (2.22), which is given by

$$(u_1+u_3\Phi_{tt}+u_4\Phi_{st}+u_5\Phi_t,u_2+u_3\Phi_{ts}+u_4\Phi_{ss}+u_5\Phi_s, u_3+u_5t_0,u_4+u_5s_0). \tag{2.31}$$

By some simple row and column transformations, we see that

$$\dim(\{(\overrightarrow{u}\cdot\overrightarrow{n_1},\overrightarrow{u}\cdot\overrightarrow{n_2},\overrightarrow{u}\cdot\overrightarrow{n_3},\overrightarrow{u}\cdot\overrightarrow{n_4}):\overrightarrow{u}\in V\}) = \dim(\{\pi_{S_1\oplus S_2}P_\xi f : f\in V\}). \tag{2.32}$$

Hence what we need to show becomes

$$\dim(V)\le \frac{5}{4}\dim(\{\pi_{S_1\oplus S_2}P_\xi f : f\in V\}) \text{ almost surely.} \tag{2.33}$$

Case $\dim(V)=1$. The is the same as the case $\dim(V)=1$ and $d=2$.

Case $\dim(V)=2$. We need to show that $\dim(\pi_{S_1\oplus S_2}P_\xi(V))=2$ almost surely. We argue by contradiction. Suppose $\dim(\pi_{S_1\oplus S_2}P_\xi(V))\le 1$ everywhere. Then

$$V=\pi_{S_1\oplus S_2}(V)\oplus S_3. \tag{2.34}$$

Let us calculate the projection of S_3 on $S_1\oplus S_2$. Take

$$f(t,s)=\Phi(t,s)=at^3+bt^2s+cts^2+ds^3. \tag{2.35}$$

Hence

$$\pi_{S_1\oplus S_2}P_\xi f=(-3at_0^2-2bt_0s_0-cs_0^2,-bt_0^2-2ct_0s_0-3ds_0^2,t_0,s_0). \tag{2.36}$$

As we know that

$$V=\pi_{S_1\oplus S_2}(V)\oplus S_3, \tag{2.37}$$

if we write $\pi_{S_1\oplus S_2}(V)=\operatorname{span}\{\overrightarrow{u}\}$ with $\overrightarrow{u}=(u_1,u_2,u_3,u_4)$, then the dimension of $\pi_{S_1\oplus S_2}P_\xi(V)$ is equal to the rank of the matrix

$$\begin{pmatrix} -3at_0^2-2bt_0s_0-cs_0^2 & -bt_0^2-2ct_0s_0-3ds_0^2 & t_0 & s_0\\ u_1 & u_2 & u_3 & u_4\end{pmatrix}. \tag{2.38}$$

For every nonzero vector $\overrightarrow{u}$, this matrix has rank two almost surely.

Case $\dim(V)=3$. We need to show that $\dim(\pi_{S_1\oplus S_2}P_\xi(V))=3$ almost surely. Suppose not. Then, by taking $\xi=(0,0)$, we obtain that $\dim(\pi_{S_1\oplus S_2}(V))=2$. Moreover,

$$V=\pi_{S_1\oplus S_2}(V)\oplus S_3. \tag{2.39}$$

Write $\pi_{S_1\oplus S_2}(V)=\operatorname{span}\{\overrightarrow{u},\overrightarrow{v}\}$ with $\overrightarrow{u}=(u_1,u_2,u_3,u_4)$ and $\overrightarrow{v}=(v_1,v_2,v_3,v_4)$. We need to show that the matrix

$$\begin{pmatrix} -3at_0^2-2bt_0s_0-cs_0^2 & -bt_0^2-2ct_0s_0-3ds_0^2 & t_0 & s_0\\ u_1 & u_2 & u_3 & u_4\\ v_1 & v_2 & v_3 & v_4\end{pmatrix} \tag{2.40}$$

has rank three almost surely. By calculating the determinants of all the 3×3 minors, it is not difficult to see that this is indeed the case.
Case $\dim(V) = 4$. We need to show that $\dim(\pi_{S_1 \oplus S_2} P_\xi(V)) = 4$ almost surely. Similar as above, we argue by contradiction. In the end, we need to show that the matrix

$$\begin{pmatrix} -3at_0^2 - 2bt_0 s_0 - cs_0^2 & -bt_0^2 - 2ct_0 s_0 - 3ds_0^2 & t_0 & s_0 \\ u_1 & u_2 & u_3 & u_4 \\ v_1 & v_2 & v_3 & v_4 \\ w_1 & w_2 & w_3 & w_4 \end{pmatrix} \tag{2.41}$$

has rank four almost surely. Suppose the determinant of the above matrix vanishes constantly. By checking the linear terms in t_0 and s_0, we obtain that

$$\det \begin{pmatrix} u_1 & u_2 & u_3 \\ v_1 & v_2 & v_3 \\ w_1 & w_2 & w_3 \end{pmatrix} = \det \begin{pmatrix} u_1 & u_2 & u_4 \\ v_1 & v_2 & v_4 \\ w_1 & w_2 & w_4 \end{pmatrix} = 0. \tag{2.42}$$

This implies that the two vectors (u_1, v_1, w_1) and (u_2, v_2, w_2) are linearly dependent. That the determinant of the matrix (2.41) vanishes constantly contradicts to the non-degeneracy of the polynomial Φ.

2.3. Multi-linear Kakeya inequalities and ball-inflation lemmas. In Proposition 2.2 we verified a transversality condition. As a consequence, we have the following multi-linear Kakeya inequality.

Lemma 2.4 (Multi-linear Kakeya)**.** *Let $d_1 = 2$ and $d_2 = 4$. For every $\iota \in \{1, 2\}$, we have the following estimate: Let $M = \Lambda K$ where Λ is the same as the one in Proposition 2.2. Let $R_1, \ldots, R_M$ be different sets from Col_K. Consider M families $\mathcal{P}_j$ consisting of rectangular boxes P in $\mathbb{R}^5$, that we refer to as plates, having the following properties:*

1. *For each $P \in \mathcal{P}_j$, there exists $\xi_j = (t_j, s_j) \in R_j$ such that d_ι sides of P have lengths equal to $R^{1/2}$ and* span $V^{(\iota)}_{\xi_j}$*, while the remaining $(5 - d_\iota)$ sides have lengths R.*
2. *All plates are subsets of a ball B_{4R} of radius $4R$.*

Then we have the following inequality:

$$\frac{1}{|B_{4R}|} \int_{B_{4R}} \Big| \prod_{j=1}^{M} F_j \Big|^{\frac{1}{M} \frac{5}{d_\iota}} \lesssim_{\epsilon,\nu} R^\epsilon \left[\prod_{j=1}^{M} \left(\frac{1}{|B_{4R}|} \Big| \int_{B_{4R}} F_j \Big| \right)^{\frac{1}{M}} \right]^{\frac{5}{d_\iota}} \tag{2.43}$$

for each function F_j of the form

$$F_j = \sum_{P \in \mathcal{P}_j} c_P 1_P. \tag{2.44}$$

The implicit constant does not depend on R or c_P.

Lemma 2.4 is essentially due to Guth [**10**], and Bennett, Bez, Flock, and Lee [**1**]. It follows from the Brascamp–Lieb inequalities in Theorem 2.1 via an induction-on-scales argument, after verifying the corresponding transversality conditions. Here we leave out the details. These multi-linear Kakeya inequalities have the following consequences.

Lemma 2.5 (Ball-inflation lemmas)**.** *Let* $d_1 = 2$ *and* $d_2 = 4$*. For every* $\iota \in \{1, 2\}$*, we have the following estimate: Let* $R_1, \ldots, R_M$ *be different squares from* Col_K*. Let* B *be an arbitrary ball in* $\mathbb{R}^5$ *of radius* $\rho^{-\iota-1}$*. Let* $\mathcal{B}$ *be a finitely overlapping cover of* B *with balls* Δ *of radius* $\rho^{-\iota}$*. For each* $g\colon [0,1]^2 \to \mathbb{C}$*, we have*

$$\frac{1}{|\mathcal{B}|}\sum_{\Delta\in\mathcal{B}}\left[\prod_{i=1}^{M}\Bigg(\sum_{\substack{R_i'\colon \textit{ square in } R_i\\ l(R_i')=\rho}}\|E_{R_i'}g\|^{\frac{d_\iota p}{5}}_{L_{\#}^{\frac{d_\iota p}{5}}(w_\Delta)}\Bigg)^{\frac{5}{d_\iota p}}\right]^{\frac{p}{M}} \tag{2.45}$$
$$\lesssim_{\epsilon,\nu}\rho^{-\epsilon}\left[\prod_{i=1}^{M}\Bigg(\sum_{\substack{R_i'\colon \textit{ square in } R_i\\ l(R_i')=\rho}}\|E_{R_i'}g\|^{\frac{d_\iota p}{5}}_{L_{\#}^{\frac{d_\iota p}{5}}(w_B)}\Bigg)^{\frac{5}{d_\iota p}}\right]^{\frac{p}{M}}.$$

Proof: We will follow the proof in Bourgain, Demeter, and Guth [**7**], that is, we will derive Lemma 2.5 from Lemma 2.4. In order to apply Lemma 2.4, we need to check that the function $|E_{R_i'}g|$ is essentially a constant on a plate whose two short sides span the linear space $V^{(1)}_{\xi_j}$ for some $\xi_j \in R_j'$. Moreover, we need to check that for each ball Δ of radius ρ^{-2}, the function $\|E_{R_i'}g\|_{L_{\#}^{\frac{4p}{5}}(w_\Delta)}$ is essentially a constant on a plate whose four short sides span the linear space $V^{(2)}_{\xi_j}$ for some $\xi_j \in R_j'$. The former statement follows from the standard Taylor expansion, and the latter one follows from Lemma 2.3. We comment here that this is what we meant previously by viewing the surface $\mathcal{S}$ as a four dimensional surface in $\mathbb{R}^5$. For the rest of the details, we refer to [**7**]. □

The idea of ball-inflations originated from the work of Bourgain, Demeter, and Guth [**7**] (see Theorem 6.6 there). If we replace the $l^{\frac{d_\iota p}{5}}$ summations over $R_i' \subset R_i$ on both sides of (2.45) by l^2 sums, essentially we arrive at the ball-inflation estimates that are proven in [**7**]. Moreover, as has been pointed out above, the proof of Theorem 6.6 in [**7**] also works for (2.45). However, there are some subtle differences when it comes to applying these ball-inflation estimates in the iteration argument in Section 5.

In [**7**], the authors there used l^2 sums over $R_i' \subset R_i$ in order to prove certain sharp l^2L^p decoupling inequalities associated to moment curves. In our case, sharp l^2L^p decouplings no long imply good enough estimates as in Theorem 1.1. Indeed, this issue already appeared in earlier attempts of trying to push the argument of [**7**] to higher dimensions, see [**6**] and [**9**] (see [**5**] for an even earlier work). In the present paper we follow the way in which ball-inflation lemmas are formulated in the work [**9**] by Zhang and the author.

3. Parabolic rescaling

In this section we state the following result which is referred to as parabolic rescaling.

Lemma 3.1. *Let $0 < \delta < \sigma \leq 1$. Then, for each square $R \subset [0,1]^2$ with side length σ and each ball $B \subset \mathbb{R}^5$ with radius δ^{-3}, we have*

$$\|E_R g\|_{L^p(w_B)} \leq B_{p,q}\Big(\frac{\delta}{\sigma}\Big)\Big(\sum_{R' \subset R: l(R')=\delta} \|E_{R'} g\|^q_{L^p(w_B)}\Big)^{1/q}. \tag{3.1}$$

The proof of this lemma is standard, see for instance Proposition 7.1 from [**5**]. One just needs to observe that our surface $\mathcal{S}$ is translation and dilation invariant, as can be seen via Lemma 2.3.

The parabolic rescaling lemma plays a significant role in decoupling theory. It is used in every iteration step. First of all, it is used to run the Bourgain–Guth scheme, in order to show the equivalence between the linear and multilinear decoupling inequalities (Theorem 4.1). Secondly, it is used in the iteration scheme in Section 5 to conclude the desired decoupling inequality (1.8).

4. Linear versus multilinear decoupling

In this section we introduce a multi-linear version of the desired decoupling inequality. Recall that K is a large number and $M = \Lambda K$. We denote by $B_{p,q}(\delta, K)$ the smallest constant such that

$$\Big\|\Big(\prod_{i=1}^{M} E_{R_i} g\Big)^{1/M}\Big\|_{L^p(w_B)} \leq B_{p,q}(\delta, K)\prod_{i=1}^{M}\Big(\sum_{R_i' \subset R_i: l(R_i')=\delta} \|E_{R_i'} g\|^q_{L^p(w_B)}\Big)^{\frac{1}{qM}} \tag{4.1}$$

holds true for all distinct squares $R_i \in \mathrm{Col}_K$, each ball $B \subset \mathbb{R}^9$ of radius δ^{-3}, and each $g\colon [0,1]^2 \to \mathbb{C}$.

By Hölder's inequality, we see that the multi-linear decoupling constant $B_{p,q}(\delta, K)$ can be controlled by the linear decoupling constant $B_{p,q}(\delta)$. It turns out that, in the case $p = q$, the reverse direction also essentially holds true. That is,

Theorem 4.1. *For each $p \geq 2$ and $K \in \mathbb{N}$, there exists $\Omega_{K,p} > 0$ and $\beta(K, p) > 0$ with*

$$\lim_{K \to \infty} \beta(K, p) = 0, \text{ for each } p, \tag{4.2}$$

such that for each small enough δ, we have

$$B_{p,p}(\delta) \leq \delta^{-\beta(K,p)-2(\frac{1}{2}-\frac{1}{p})} + \Omega_{K,p} \log_K \left(\frac{1}{\delta}\right) \max_{\delta \leq \delta' \leq 1} \left(\frac{\delta'}{\delta}\right)^{2(\frac{1}{2}-\frac{1}{p})} B_{p,p}(\delta', K). \tag{4.3}$$

The proof of this theorem is standard, and is essentially the same as that of Theorem 8.1 from [**5**]. Hence we will only give a sketch of the proof.

Proof of Theorem 4.1. The proof is via the Bourgain–Guth scheme [**8**]. For a given $\alpha \in \mathrm{Col}_K$ and a ball $B' \subset \mathbb{R}^5$ of radius K, by the uncertainty principle, we know that $|E_\alpha g|$ is essentially constant on B'. This suggests us to define

$$c_\alpha := \|E_\alpha g\|_{L^p_\#(B')}. \tag{4.4}$$

We temporarily fix B'. Denote by α^* the cube that maximises $\{c_{\alpha'} : \alpha' \in \mathrm{Col}_K\}$. Let Col_K^* be the collection of those cubes α' with

$$c_{\alpha'} \geq K^{-20} c_{\alpha^*}. \tag{4.5}$$

There are no particular reasons why we picked K^{-20}. It can also be K^{-30} or something that is even smaller.

Next we analyse Col_K^*. There are two cases. A first case is that the cardinality of Col_K^* is smaller than ΛK. Here Λ is the same constant as in Lemma 2.4. A second case is of course that the cardinality of Col_K^* is bigger than or equal to ΛK.

In the first case, we have

$$\|E_{[0,1]^2} g\|_{L^p(w_{B'})} \leq 2\|E_{\alpha^*} g\|_{L^p(w_{B'})} + \Big\| \sum_{\alpha' \in \mathrm{Col}_K^*} E_{\alpha'} g \Big\|_{L^p(w_{B'})}. \tag{4.6}$$

By L^2 orthogonality, the right hand side can be bounded by

$$K^{2(\frac{1}{2}-\frac{1}{p})}\Big(\sum_{\alpha'\in \mathrm{Col}_K^*}\|E_{\alpha'}g\|^p_{L^p(w_{B'})}\Big)^{1/p} \tag{4.7}$$
$$\leq K^{2(\frac{1}{2}-\frac{1}{p})}\Big(\sum_{\alpha\in \mathrm{Col}_K}\|E_{\alpha}g\|^p_{L^p(w_{B'})}\Big)^{1/p}.$$

In the second case, we are able to find $R_1,\dots,R_{\Lambda K}\in \mathrm{Col}_K$ such that

$$\|E_{[0,1]^2}g\|_{L^p(w_{B'})}\leq K^{20}\Big\|\prod_{j=1}^{\Lambda K}|E_{R_j}g|^{\frac{1}{\Lambda K}}\Big\|_{L^p(w_{B'})}. \tag{4.8}$$

In either case we have

$$\|E_{[0,1]^2}g\|_{L^p(w_{B'})}\leq K^{2(\frac{1}{2}-\frac{1}{p})}\Big(\sum_{\alpha\in \mathrm{Col}_K}\|E_{\alpha}g\|^p_{L^p(w_{B'})}\Big)^{1/p} \tag{4.9}$$
$$+\max_{R_1,\dots,R_{\Lambda K}}K^{20}\Big\|\prod_{j=1}^{\Lambda K}|E_{R_j}g|^{\frac{1}{\Lambda K}}\Big\|_{L^p(w_{B'})}.$$

We raise both sides to the p-th power and sum over balls B' inside a ball B of radius δ^{-3}. In the end, we obtain

$$\|E_{[0,1]^2}g\|_{L^p(w_{B})}\leq K^{2(\frac{1}{2}-\frac{1}{p})}\Big(\sum_{\alpha\in \mathrm{Col}_K}\|E_{\alpha}g\|^p_{L^p(w_{B})}\Big)^{1/p} \tag{4.10}$$
$$+\sum_{R_1,\dots,R_{\Lambda K}}K^{20}\Big\|\prod_{j=1}^{\Lambda K}|E_{R_j}g|^{\frac{1}{\Lambda K}}\Big\|_{L^p(w_{B})}.$$

To the last term in the above display, we apply the definition of the multilinear decoupling constant. This leads to an estimate

$$\|E_{[0,1]^2}g\|_{L^p(w_{B})}\leq K^{2(\frac{1}{2}-\frac{1}{p})}\Big(\sum_{\alpha\in \mathrm{Col}_K}\|E_{\alpha}g\|^p_{L^p(w_{B})}\Big)^{1/p} \tag{4.11}$$
$$+C_K B_{p,p}(\delta,K)\Big(\sum_{l(\Delta)=\delta}\|E_{\Delta}g\|^p_{L^p(w_{B})}\Big)^{1/p}.$$

Here C_K is a large constant depending only on K. Its size is not relevant to us. However, having a correct power of K in the first term on the right hand side of (4.11) is crucial. We will iterate (4.11) with $E_\alpha g$ in place of $E_{[0,1]^2}g$ on the left hand side, until we reach the frequency scale δ. In other words, we will iterate (4.11) for $\log_K\left(\frac{1}{\delta}\right)$ many times. This will result in the desired estimate (4.3). In particular, the main power of δ^{-1} in the first term on the right hand side of (4.3), which is $\frac{1}{2}-\frac{1}{p}$, is exactly inherited from the power of K in the first term on the right hand side of (4.11). Moreover, the constant $\Omega_{K,p}$ depends only on K and p, and it will not play any particularly important role later. We leave out the details of this step of iteration. □

5. Iteration

In this section we run the final iteration argument. The consequence of this iteration, combined with Theorem 4.1, will lead to the desired decoupling inequality (1.8).

There will be two terms that are involved in the iteration procedure. They are

$$D_p(q,B^r) := \left(\prod_{i=1}^{M}\sum_{J_{i,q}\subset R_i}\|E_{J_{i,q}}g\|^p_{L^p_{\#}(w_{B^r})}\right)^{\frac{1}{pM}} \tag{5.1}$$

and

$$A_p(q,B^r,s) = \left(\frac{1}{|\mathcal{B}_s(B^r)|}\sum_{B^s\in\mathcal{B}_s(B^r)} D_2(q,B^s)^p\right)^{1/p}. \tag{5.2}$$

Here, for a positive number r, we use B^r to denote a ball of radius δ^{-r} and $\mathcal{B}_s(B^r)$ denotes a finitely overlapping collection of balls B^s that lie inside of a ball B^r. In the notation $J_{i,q}$, the index i indicates that this square lies in R_i and q indicates that the square $J_{i,q}$ has side length δ^q.

Define $\alpha_1,\alpha_2,\beta_2\in(0,1)$ as follows

$$\frac{1}{\frac{2p}{5}} = \frac{\alpha_1}{\frac{4p}{5}} + \frac{1-\alpha_1}{2},$$

$$\frac{1}{\frac{4p}{5}} = \frac{\alpha_2}{p} + \frac{1-\alpha_2}{6},$$

$$\frac{1}{6} = \frac{1-\beta_2}{2} + \frac{\beta_2}{\frac{4p}{5}}.$$

We will start our iteration with the term

$$A_p(1,B^3,1) = \left(\frac{1}{|\mathcal{B}_1(B^3)|}\sum_{B^1\in\mathcal{B}_1(B^3)} D_2(1,B^1)^p\right)^{1/p}. \tag{5.3}$$

By Hölder's inequality, it can be bounded by

$$\delta^{-2(\frac{1}{2}-\frac{5}{2p})}\Bigg(\frac{1}{|\mathcal{B}_1(B^3)|}\sum_{B^1\in\mathcal{B}_1(B^3)} D_{\frac{2p}{5}}(1,B^1)^p\Bigg)^{1/p}. \tag{5.4}$$

We apply Lemma 2.5 with $\iota=1$ to (5.4) and bound it by

$$\delta^{-2(\frac{1}{2}-\frac{5}{2p})-\epsilon}\Bigg(\frac{1}{|\mathcal{B}_2(B^3)|}\sum_{B^2\in\mathcal{B}_2(B^3)} D_{\frac{2p}{5}}(1,B^2)^p\Bigg)^{1/p}. \tag{5.5}$$

By Hölder's inequality, the right hand side of (5.5) can be dominated by

$$\begin{aligned}\delta^{-2(\frac{1}{2}-\frac{5}{2p})-\epsilon}&\Bigg(\frac{1}{|\mathcal{B}_2(B^3)|}\sum_{B^2\in\mathcal{B}_2(B^3)} D_{\frac{4p}{5}}(1,B^2)^p\Bigg)^{\frac{\alpha_1}{p}}\\ &\times\Bigg(\frac{1}{|\mathcal{B}_2(B^3)|}\sum_{B^2\in\mathcal{B}_2(B^3)} D_2(1,B^2)^p\Bigg)^{\frac{1-\alpha_1}{p}}.\end{aligned} \tag{5.6}$$

By L^2 orthogonality, this can be bounded by

$$\delta^{-2(\frac{1}{2}-\frac{5}{2p})-\epsilon}\Bigg(\frac{1}{|\mathcal{B}_2(B^3)|}\sum_{B^2\in\mathcal{B}_2(B^3)} D_{\frac{4p}{5}}(1,B^2)^p\Bigg)^{\frac{\alpha_1}{p}} A_p(2,B^3,2)^{1-\alpha_1}. \tag{5.7}$$

In the next step, we apply Lemma 2.5 with $\iota=2$ and obtain

$$\delta^{-2(\frac{1}{2}-\frac{5}{2p})-\epsilon}D_{\frac{4p}{5}}(1,B^3)^{\alpha_1}A_p(2,B^3,2)^{1-\alpha_1}. \tag{5.8}$$

The last term $A_p(2,B^3,2)^{1-\alpha_1}$ is ready for iteration. We further process the D-term. By Hölder's inequality

$$D_{\frac{4p}{5}}(1,B^3)\lesssim D_6(1,B^3)^{1-\alpha_2}D_p(1,B^3)^{\alpha_2}. \tag{5.9}$$

The second term on the right hand side is already of the form of the term in the decoupling inequality (1.7) and it will not be further processed. It is the former term on the right hand side that needs further process.

Notice that in the term $D_6(1,B^3)$, we are dealing with terms $\|E_{J_{i,1}}g\|_{L^6_{\#}(w_{B^3})}$. By the uncertainty principle, such a ball of radius δ^{-3} is not able to distinguish the surface $\mathcal{S}$ from

$$\{(t,s,\Phi_t(t,s),\Phi_s(t,s),0):(t,s)\in J_{i,1}\} \tag{5.10}$$

under certain affine transformations. By the l^6L^6 decoupling estimate for the surface (5.10) obtained in [**4**], we obtain

$$D_6(1,B^3)\lesssim \delta^{-(\frac{1}{2}-\frac{1}{6})-\epsilon}D_6\Big(\frac{3}{2},B^3\Big). \tag{5.11}$$

By Hölder's inequality, this can be further bounded by

$$
\begin{aligned}
D_6(1,B^3) &\lesssim \delta^{-(\frac{1}{2}-\frac{1}{6})-\epsilon} D_6\left(\frac{3}{2},B^3\right)\\
&\lesssim \delta^{-(\frac{1}{2}-\frac{1}{6})-\epsilon} D_2\left(\frac{3}{2},B^3\right)^{1-\beta_2} D_{\frac{4p}{5}}\left(\frac{3}{2},B^3\right)^{\beta_2} \qquad (5.12)\\
&\lesssim \delta^{-(\frac{1}{2}-\frac{1}{6})-\epsilon} D_2(3,B^3)^{1-\beta_2} D_{\frac{4p}{5}}\left(\frac{3}{2},B^3\right)^{\beta_2}.
\end{aligned}
$$

In the last step, we applied L^2 orthogonality. In the end, what we have obtained so far can be organised as

$$
\begin{aligned}
A_p(1,B^3,1) \lesssim_{\epsilon,K} & \ \delta^{-\epsilon-2(\frac{1}{2}-\frac{5}{2p})-(\frac{1}{2}-\frac{1}{6})\alpha_1(1-\alpha_2)}\\
&\times A_p(2,B^3,2)^{1-\alpha_1} A_p(3,B^3,3)^{\alpha_1(1-\alpha_2)(1-\beta_2)} \qquad (5.13)\\
&\times D_{\frac{4p}{5}}\left(\frac{3}{2},B^3\right)^{\alpha_1(1-\alpha_2)\beta_2} D_p(1,B^3)^{\alpha_1\alpha_2}.
\end{aligned}
$$

Now we run this iteration procedure for r many times. For all balls B of radius $\delta^{-2\cdot(\frac{3}{2})^r}$, we have

(5.14)

$$
\begin{aligned}
A_p(1,B,1) \lesssim_{\epsilon,r,K} & \left(\frac{1}{\delta}\right)^{\epsilon+2(\frac{1}{2}-\frac{5}{2p})} \times \underbrace{\prod_{i=0}^{r-1}\left(\frac{1}{\delta}\right)^{(\frac{3}{2})^i(\frac{1}{2}-\frac{1}{6})\alpha_1(1-\alpha_2)[(1-\alpha_2)\beta_2]^i}}_{l^6L^6 \text{ decoupling}}\\
&\times A_p(2,B,2)^{1-\alpha_1} D_{\frac{4p}{5}}\left(\left(\frac{3}{2}\right)^r,B\right)^{\alpha_1[(1-\alpha_2)\beta_2]^r}\\
&\times \left(\prod_{i=1}^{r} A_p\left(2\left(\frac{3}{2}\right)^i,B,2\left(\frac{3}{2}\right)^i\right)^{\alpha_1(1-\alpha_2)(1-\beta_2)[(1-\alpha_2)\beta_2]^{i-1}}\right)\\
&\times \left(\prod_{i=0}^{r-1} D_p\left(\left(\frac{3}{2}\right)^i,B\right)^{\alpha_1\alpha_2[(1-\alpha_2)\beta_2]^i}\right).
\end{aligned}
$$

Define

(5.15)

$$\gamma_0 = 1-\alpha_1;\ \gamma_i = \alpha_1(1-\alpha_2)(1-\beta_2)[(1-\alpha_2)\beta_2]^{i-1},\ \text{for } 1\le i\le r;$$

$$b_i = 2\cdot\left(\frac{3}{2}\right)^i,\ \text{for } 0\le i\le r;$$

$$\tau_r = \alpha_1[(1-\alpha_2)\beta_2]^r;\ \tau_i = \alpha_1\alpha_2[(1-\alpha_2)\beta_2]^i,\ \text{for } 0\le i\le r-1;$$

$$w_i = \frac{1-\alpha_2}{2\alpha_2}\tau_i,\ \text{for } 0\le i\le r-1.$$

We can write, using Hölder's inequality,

$$D_{\frac{4p}{5}}\left(\left(\frac{3}{2}\right)^r, B\right) \lesssim \left(\frac{1}{\delta}\right)^{2(\frac{3}{2})^r\cdot\frac{1}{4p}} D_p\left(\left(\frac{3}{2}\right)^r, B\right).$$

With these, the estimate (5.14) becomes

$$\begin{aligned} A_p(1,B,1) \lesssim_{r,\epsilon,K} &\left(\frac{1}{\delta}\right)^{\epsilon+2(\frac{1}{2}-\frac{5}{2p})} \\ &\times \underbrace{\left(\frac{1}{\delta}\right)^{2(\frac{3}{2})^r\cdot\frac{1}{4p}\cdot\alpha_1[(1-\alpha_2)\beta_2]^r}}_{(\star)} \left(\prod_{i=0}^{r-1}\left(\frac{1}{\delta}\right)^{(\frac{1}{2}-\frac{1}{6})b_iw_i}\right) \\ &\times \left(\prod_{i=0}^{r} A_p(b_i,B,b_i)^{\gamma_i}\right)\left(\prod_{i=0}^{r} D_p\left(\frac{b_i}{2},B\right)^{\tau_i}\right). \end{aligned} \tag{5.16}$$

When p is close to 9, by a simple calculation, we see that $3(1-\alpha_2)\beta_2/2 < 1$. Hence the contribution from the term $(\star)$ can be absorbed by $\delta^{-\epsilon}$ when r is chosen to be large enough. Using this and a simple rescaling argument, we can rewrite (5.14) as follows

$$\begin{aligned} A_p(u,B,u) \lesssim_{r,\epsilon,K} &\left(\frac{1}{\delta}\right)^{\epsilon+2u(\frac{1}{2}-\frac{5}{2p})}\left(\prod_{i=0}^{r-1}\left(\frac{1}{\delta}\right)^{(\frac{1}{2}-\frac{1}{6})ub_iw_i}\right) \\ &\times \left(\prod_{i=0}^{r} A_p(b_iu,B,b_iu)^{\gamma_i}\right)\left(\prod_{i=0}^{r} D_p\left(\frac{b_iu}{2},B\right)^{\tau_i}\right). \end{aligned} \tag{5.17}$$

Here B stands for a ball of radius δ^{-3}, and u is a sufficiently small positive constant such that $u\cdot(\frac{3}{2})^r \le 1$.

In the end, we iterate (5.17). To iterate, we will dominate each $A_p(ub_i, B, ub_i)$ again by using (5.17). To enable such an iteration, we need to choose u to be even smaller. Let M be a large integer. Choose u such that

$$\left[2\left(\frac{3}{2}\right)^r\right]^M u \leq 2. \tag{5.18}$$

This allows us to iterate (5.17) M times. To simplify the iteration, we bound all the powers of $\frac{1}{\delta}$ by

$$2u\left(\frac{1}{2}-\frac{5}{2p}\right)+\left(\sum_{i=0}^{\infty} u\left(\frac{1}{2}-\frac{1}{6}\right)b_i w_i\right). \tag{5.19}$$

By a direct calculation,

$$\sum_{j=0}^{\infty} b_j w_j = \frac{3(-5+p)}{2(15-10p+p^2)}. \tag{5.20}$$

Moreover,

$$\sum_{j=0}^{\infty} b_j \tau_j = \frac{75-25p+2p^2}{15-10p+p^2}. \tag{5.21}$$

If we define

$$\lambda_0 := 2\left(\frac{1}{2}-\frac{5}{2p}\right)+\left(\frac{1}{2}-\frac{1}{6}\right)\frac{3(-5+p)}{2(15-10p+p^2)}, \tag{5.22}$$

then (5.17) can be rewritten as follows

$$\begin{aligned} A_p(u,B,u) &\lesssim_{r,\epsilon,K} \delta^{-\epsilon-u\lambda_0} \\ &\times \left(\prod_{i=0}^{r} A_p(ub_i,B,ub_i)^{\gamma_i}\right)\left(\prod_{i=0}^{r} D_p\left(\frac{ub_i}{2},B\right)^{\tau_i}\right), \end{aligned} \tag{5.23}$$

for every ball B of radius δ^{-3}. Notice that the implicit constant here is allowed to depend on all the parameters ϵ, r, and K. Now we have arrived precisely at the estimate (6.51) from [**6**]. The calculation there, from page 27 to page 30, can be repeated line by line. In the end, we obtain that

$$\log_{\frac{1}{\delta}} B_{9,9}(\delta) \leq \lim_{p\to 9} \frac{\lambda_0(p)}{\frac{1}{2}\left(\sum_{j=0}^{\infty} b_j\tau_j(p)\right)}. \tag{5.24}$$

By plugging in the calculation (5.20)–(5.21), we will be able to conclude the desired decoupling inequality (1.8).

For the sake of completeness, we include the proof of (5.24). We iterate the above estimate (5.23) M times, and obtain

(5.25)

$$A_p(u,B,u) \lesssim_{\epsilon,r,K,M} \delta^{-u\lambda_0-\epsilon}\left(\prod_{j_1=0}^{r}\delta^{-u\lambda_0 b_{j_1}\gamma_{j_1}}\right)\times\cdots$$
$$\times\left(\prod_{j_1=0}^{r}\prod_{j_2=0}^{r}\cdots\prod_{j_{M-1}=0}^{r}\delta^{-u\lambda_0 b_{j_1}b_{j_2}\dots b_{j_{M-1}}\gamma_{j_1}\gamma_{j_2}\dots\gamma_{j_{M-1}}}\right)$$
$$\times\left(\prod_{j_1=0}^{r}D_p\left(\frac{u}{2}\cdot b_{j_1},B\right)^{\tau_{j_1}}\right)\left(\prod_{j_1=0}^{r}\prod_{j_2=0}^{r}D_p\left(\frac{u}{2}\cdot b_{j_1}b_{j_2},B\right)^{\tau_{j_1}\gamma_{j_2}}\right)\times\cdots$$
$$\times\left(\prod_{j_1=0}^{r}\prod_{j_2=0}^{r}\cdots\prod_{j_M=0}^{r}D_p\left(\frac{u}{2}\cdot b_{j_1}b_{j_2}\dots b_{j_M},B\right)^{\tau_{j_1}\gamma_{j_2}\dots\gamma_{j_M}}\right)$$
$$\times\left(\prod_{j_1=0}^{r}\prod_{j_2=0}^{r}\cdots\prod_{j_M=0}^{r}A_p(u\cdot b_{j_1}b_{j_2}\dots b_{j_M},B,u\cdot b_{j_1}b_{j_2}\dots b_{j_M})^{\gamma_{j_1}\gamma_{j_2}\dots\gamma_{j_M}}\right).$$

Collecting the powers of $\frac{1}{\delta}$. We obtain

$$u\lambda_0+u\lambda_0\left(\sum_{j=0}^{r}b_j\gamma_j\right)+\cdots+u\lambda_0\left(\sum_{j=0}^{r}b_j\gamma_j\right)^{M-1} = u\lambda_0\cdot\frac{1-(\sum_{j=0}^{r}b_j\gamma_j)^M}{1-(\sum_{j=0}^{r}b_j\gamma_j)}. \tag{5.26}$$

The contribution from the D_p-terms. By parabolic rescaling (Lemma 3.1), the product of all these D_p-terms can be controlled by

(5.27)

$$\left(\prod_{j_1=0}^{r}B_{p,p}(\delta^{1-\frac{u}{2}b_{j_1}})^{\tau_{j_1}}D_p(1,B)^{\tau_{j_1}}\right)$$
$$\times\left(\prod_{j_1=0}^{r}\prod_{j_2=0}^{r}B_{p,p}(\delta^{1-\frac{u}{2}b_{j_1}b_{j_2}})^{\tau_{j_1}\gamma_{j_2}}D_p(1,B)^{\tau_{j_1}\gamma_{j_2}}\right)\times\cdots$$
$$\times\left(\prod_{j_1=0}^{r}\prod_{j_2=0}^{r}\cdots\prod_{j_M=0}^{r}B_{p,p}(\delta^{1-\frac{u}{2}b_{j_1}b_{j_2}\dots b_{j_M}})^{\tau_{j_1}\gamma_{j_2}\dots\gamma_{j_M}}D_p(1,B)^{\tau_{j_1}\gamma_{j_2}\dots\gamma_{j_M}}\right)$$
$$\lesssim\left(\prod_{j_1=0}^{r}B_{p,p}(\delta^{1-\frac{u}{2}b_{j_1}})^{\tau_{j_1}}\right)\times\left(\prod_{j_1=0}^{r}\prod_{j_2=0}^{r}B_{p,p}(\delta^{1-\frac{u}{2}b_{j_1}b_{j_2}})^{\tau_{j_1}\gamma_{j_2}}\right)\times\cdots$$
$$\times\left(\prod_{j_1=0}^{r}\prod_{j_2=0}^{r}\cdots\prod_{j_M=0}^{r}B_{p,p}(\delta^{1-\frac{u}{2}b_{j_1}b_{j_2}\dots b_{j_M}})^{\tau_{j_1}\gamma_{j_2}\dots\gamma_{j_M}}\right)$$
$$\times(D_p(1,B))^{1-(\sum_{j=0}^{r}\gamma_j)^M}.$$

The contribution from the A_p-term. By Hölder's inequality, this term can be bounded by

$$\prod_{j_1=0}^{r}\cdots\prod_{j_M=0}^{r}\left(\frac{1}{\delta}\right)^{2ub_{j_1}\dots b_{j_M}\gamma_{j_1}\dots\gamma_{j_M}}[D_p(b_{j_1}\dots b_{j_M}u,B)]^{\gamma_{j_1}\dots\gamma_{j_M}}. \tag{5.28}$$

To control the D_p term, we again invoke the parabolic rescaling and bound the last expression by

$$\left(\frac{1}{\delta}\right)^{2u(\sum_{j=0}^{r}b_j\gamma_j)^M} \times\prod_{j_1=0}^{r}\cdots\prod_{j_M=0}^{r}(B_{p,p}(\delta^{1-ub_{j_1}\dots b_{j_M}}))^{\gamma_{j_1}\dots\gamma_{j_M}}(D_p(1,B))^{\gamma_{j_1}\dots\gamma_{j_M}}. \tag{5.29}$$

So far we have obtained

Proposition 5.1. *For each ball B of radius δ^{-3}, and for each sufficiently small u, we have*

$$\begin{aligned} A_p(u,B,u) \lesssim_{\epsilon,r,M,K} &\left(\frac{1}{\delta}\right)^{\epsilon+u\lambda_0\cdot\frac{1-(\sum_{j=0}^{r}b_j\gamma_j)^M}{1-(\sum_{j=0}^{r}b_j\gamma_j)}+2u(\sum_{j=0}^{r}b_j\gamma_j)^M} D_p(1,B)\\ &\times\left(\prod_{j_1=0}^{r}B_{p,p}(\delta^{1-\frac{u}{2}b_{j_1}})^{\tau_{j_1}}\right)\\ &\times\left(\prod_{j_i=0}^{r}\prod_{j_2=0}^{r}B_{p,p}(\delta^{1-\frac{u}{2}b_{j_1}b_{j_2}})^{\tau_{j_1}\gamma_{j_2}}\right)\times\cdots\\ &\times\left(\prod_{j_1=0}^{r}\prod_{j_2=0}^{r}\cdots\prod_{j_M=0}^{r}B_{p,p}(\delta^{1-\frac{u}{2}b_{j_1}b_{j_2}\dots b_{j_M}})^{\tau_{j_1}\gamma_{j_2}\dots\gamma_{j_M}}\right)\\ &\times\left(\prod_{j_1=0}^{r}\cdots\prod_{j_M=0}^{r}(B_{p,p}(\delta^{1-ub_{j_1}\dots b_{j_M}}))^{\gamma_{j_1}\dots\gamma_{j_M}}\right). \end{aligned} \tag{5.30}$$

The final step of the proof. Now we come to the final step of the proof for the desired estimate (1.8) at the critical exponent $p = 9$. We will combine Theorem 4.1 with Proposition 5.1. Let η_p be the unique number such that

$$\lim_{\delta\to 0}\frac{B_{p,p}(\delta)}{\delta^{-(\eta_p+\mu)}} = 0,\ \text{for each } \mu > 0, \tag{5.31}$$

and

$$\limsup_{\delta\to 0}\frac{B_{p,p}(\delta)}{\delta^{-(\eta_p-\mu)}} = \infty,\ \text{for each } \mu > 0. \tag{5.32}$$

Let B have radius δ^{-3}. We substitute the bound $B_{p,p}(\delta) \lesssim_\mu \delta^{-(\eta_p+\mu)}$ into the right hand side of (5.30), and obtain

$$A_p(u, B, u) \lesssim_{r,M,K,\mu} \delta^{-\eta_{p,\mu,u,r,M}} D_p(1, B), \tag{5.33}$$

where

$$\eta_{p,\mu,u,r,M} = u\lambda_0 \cdot \frac{1-(\sum_{j=0}^r b_j\gamma_j)^M}{1-(\sum_{j=0}^r b_j\gamma_j)} + 2u\Big(\sum_{j=0}^r b_j\gamma_j\Big)^M + (\mu+\eta_p)\left[1-u\cdot\Big(\sum_{j=0}^r b_j\gamma_j\Big)^M - \frac{u}{2}\Big(\sum_{j=0}^r b_j\tau_j\Big)\frac{1-(\sum_{j=0}^r b_j\gamma_j)^M}{1-(\sum_{j=0}^r b_j\gamma_j)}\right]. \tag{5.34}$$

This, combined with the Cauchy–Schwarz inequality, implies

$$\left\|\Big(\prod_{i=1}^{\Lambda K} E_{R_i} g\Big)^{\frac{1}{\Lambda K}}\right\|_{L^p_\#(w_B)} \lesssim_{r,M,K,\mu} \delta^{-100u-\eta_{p,\mu,u,r,M}} D_p(1, B). \tag{5.35}$$

Here 100 is some large number that can certainly be replaced by something much smaller. However, as u is arbitrarily small, we can afford losing such a large constant. By taking the supremum over g and R_i (with fixed K) in the above estimate, we obtain

$$B_{p,p}(\delta, K) \lesssim_{r,M,K,\mu} \delta^{-\tilde{\eta}_{p,\mu,u,r,M}}, \tag{5.36}$$

where

$$\tilde{\eta}_{p,\mu,u,r,M} := \eta_{p,\mu,u,r,M} + 100u. \tag{5.37}$$

We move η_p from the right hand side of the expression (5.34) to the left hand side, and then divide both sides by u to obtain

$$\frac{1}{u}(\tilde{\eta}_{p,\mu,u,r,M}-\eta_p) = 100 + \frac{\mu}{u} + \lambda_0 \cdot \frac{1-(\sum_{j=0}^r b_j\gamma_j)^M}{1-(\sum_{j=0}^r b_j\gamma_j)} + 2\Big(\sum_{j=0}^r b_j\gamma_j\Big)^M - (\mu+\eta_p)\left[\Big(\sum_{j=0}^r b_j\gamma_j\Big)^M + \frac{1}{2}\Big(\sum_{j=0}^r b_j\tau_j\Big)\frac{1-(\sum_{j=0}^r b_j\gamma_j)^M}{1-(\sum_{j=0}^r b_j\gamma_j)}\right]. \tag{5.38}$$

Our goal is to show that

$$\eta_9 \le 2\Big(\frac{1}{2}-\frac{1}{9}\Big). \tag{5.39}$$

We argue by contradiction. Suppose that

$$\eta_9 > \frac{7}{9}. \tag{5.40}$$

Then, for sufficiently small $\epsilon_1 > 0$ this forces

$$\eta_p > \frac{7}{9}, \text{ for each } p \in (9-\epsilon_1, 9). \tag{5.41}$$

We rewrite the right hand side of (5.38) as

$$\underbrace{\left(\lambda_0 - \frac{1}{2}\cdot(\mu+\eta_p)\left(\sum_{j=0}^{r} b_j\tau_j\right)\right)\frac{1-(\sum_{j=0}^{r} b_j\gamma_j)^M}{1-(\sum_{j=0}^{r} b_j\gamma_j)}}_{(\star)} + 100 + \frac{\mu}{u} + (2-\mu-\eta_p)\left(\sum_{j=0}^{r} b_j\gamma_j\right)^M. \tag{5.42}$$

It transpires that the term $(\star)$ is dominant. Indeed, by a direct calculation, when p is smaller than (and sufficiently close to) the critical exponent 9, we have

$$\sum_{j=0}^{\infty} b_j\gamma_j > 1. \tag{5.43}$$

Moreover,

$$\lim_{p\to 9}\sum_{j=0}^{\infty} b_j\gamma_j = 1. \tag{5.44}$$

In addition to these, by a direct calculation we observe that

$$\lim_{p\to 9}\frac{\lambda_0}{\frac{1}{2}\cdot(\sum_{j=0}^{\infty} b_j\tau_j)} = \frac{7}{9}. \tag{5.45}$$

Choose now p close enough to 9, r and M large enough, and then μ small enough. By combining (5.41), (5.43), (5.44), and (5.45) we obtain that for these values of p, r, M, and μ, the expression appearing in (5.42) is negative. Going back to (5.38), for these values of p, μ, r, and M we conclude that

$$\tilde{\eta}_{p,\mu,u,r,M} < \eta_p. \tag{5.46}$$

Together with (5.41), for an exponent p slightly smaller than the critical exponent 9 and for K large enough, Theorem 4.1 implies that

$$B_{p,p}(\delta) \le \Omega_{K,p}\log_K\left(\frac{1}{\delta}\right)\max_{\delta\le\delta'\le 1}\left(\frac{\delta'}{\delta}\right)^{2(\frac{1}{2}-\frac{1}{p})} B_{p,p}(\delta', K). \tag{5.47}$$

We have two possibilities. First, if

$$\tilde{\eta}_{p,\mu,u,r,M} < 2\left(\frac{1}{2} - \frac{1}{p}\right), \tag{5.48}$$

then (5.47) combined with (5.36) forces

$$B_{p,p}(\delta) \lesssim_\epsilon \left(\frac{1}{\delta}\right)^{\epsilon+2(\frac{1}{2}-\frac{1}{p})}.$$

This contradicts (5.41).

Second, if

$$\tilde{\eta}_{p,\mu,u,r,M} \geq 2\left(\frac{1}{2} - \frac{1}{p}\right), \tag{5.49}$$

then again (5.47) combined with (5.36) forces

$$B_{p,p}(\delta) \lesssim_\epsilon \left(\frac{1}{\delta}\right)^{\epsilon+\tilde{\eta}_{p,\mu,u,r,M}}.$$

This contradicts (5.46). Since both cases lead to a contradiction, it can only be that our original assumption (5.40) is false. This finishes the proof of (5.39).

References

[1] J. Bennett, N. Bez, T. C. Flock, and S. Lee, Stability of the Brascamp–Lieb constant and applications, *Amer. J. Math.* **140(2)** (2018), 543–569. DOI: 10.1353/ajm.2018.0013.

[2] J. Bennett, A. Carbery, M. Christ, and T. Tao, Finite bounds for Hölder–Brascamp–Lieb multilinear inequalities, *Math. Res. Lett.* **17(4)** (2010), 647–666. DOI: 10.4310/MRL.2010.v17.n4.a6.

[3] J. Bourgain and C. Demeter, The proof of the l^2 decoupling conjecture, *Ann. of Math. (2)* **182(1)** (2015), 351–389. DOI: 10.4007/annals.2015.182.1.9.

[4] J. Bourgain and C. Demeter, Decouplings for surfaces in $\mathbb{R}^4$, *J. Funct. Anal.* **270(4)** (2016), 1299–1318. DOI: 10.1016/j.jfa.2015.11.008.

[5] J. Bourgain and C. Demeter, Mean value estimates for Weyl sums in two dimensions, *J. Lond. Math. Soc. (2)* **94(3)** (2016), 814–838. DOI: 10.1112/jlms/jdw063.

[6] J. Bourgain, C. Demeter, and S. Guo, Sharp bounds for the cubic Parsell–Vinogradov system in two dimensions, *Adv. Math.* **320** (2017), 827–875. DOI: 10.1016/j.aim.2017.09.008.

[7] J. Bourgain, C. Demeter, and L. Guth, Proof of the main conjecture in Vinogradov's Mean Value Theorem for degrees higher than three, *Ann. of Math. (2)* **184(2)** (2016), 633–682. DOI: 10.4007/annals.2016.184.2.7.

[8] J. Bourgain and L. Guth, Bounds on oscillatory integral operators based on multilinear estimates, *Geom. Funct. Anal.* **21(6)** (2011), 1239–1295. DOI: 10.1007/s00039-011-0140-9.

[9] S. Guo and R. Zhang, On integer solutions of Parsell–Vinogradov systems, *Invent. Math.* **218(1)** (2019), 1–81. DOI: 10.1007/s00222-019-00881-6.

[10] L. Guth, A short proof of the multilinear Kakeya inequality, *Math. Proc. Cambridge Philos. Soc.* **158(1)** (2015), 147–153. DOI: 10.1017/S0305004114000589.

[11] K. Henriot, Additive equations in dense variables via truncated restriction estimates, *Proc. Lond. Math. Soc. (3)* **114(5)** (2017), 927–959. DOI: 10.1112/plms.12028.

[12] S. T. Parsell, S. M. Prendiville, and T. D. Wooley, Near-optimal mean value estimates for multidimensional Weyl sums, *Geom. Funct. Anal.* **23(6)** (2013), 1962–2024. DOI: 10.1007/s00039-013-0242-7.

[13] S. M. Prendiville, Solution-free sets for sums of binary forms, *Proc. Lond. Math. Soc. (3)* **107(2)** (2013), 267–302. DOI: 10.1112/plms/pds083.

[14] Y. Tschinkel, Algebraic varieties with many rational points, in: *"Arithmetic Geometry"*, Clay Math. Proc. **8**, Amer. Math. Soc., Providence, RI, 2009, pp. 243–334.

[15] K. Van Valckenborgh, Squareful points of bounded height, *C. R. Math. Acad. Sci. Paris* **349(11–12)** (2011), 603–606. DOI: 10.1016/j.crma.2011.05.001.

[16] R. Wongkew, Volumes of tubular neighbourhoods of real algebraic varieties, *Pacific J. Math.* **159(1)** (1993), 177–184. DOI: 10.2140/pjm.1993.159.177.

Department of Mathematics, Indiana University, 831 East 3rd St., Bloomington IN 47405

E-mail address: shaomingguo2018@gmail.com

Primera versió rebuda el 20 d'abril de 2018,
darrera versió rebuda el 24 de maig de 2019.